Chapter 6

Assimilation Under Ideal Conditions

In this Chapter, we present the results from a series of synthetic experiments under ideal assimilation conditions. By ideal conditions we mean that the same hydrologic model, the same model inputs, and the same error statistics are used for the generation of the (synthetic) true fields and for the assimilation. Experiments under more realistic nonideal conditions are presented in Chapter 7.

Even though the estimates derived in ideal experiments are unrealistically good, such experiments are valuable. First, assimilation under ideal conditions allows us to test whether the estimation algorithm is working properly. If conditions were not ideal, any coding error can be equally well blamed on the nonideal nature of the assimilation, and we would be left with an inconclusive result. Second, it is much easier to portray and to understand the basic features of the assimilation algorithm when conditions are ideal. Finally, we can gain valuable information on the sensitivity of the estimation procedure to various inputs.

In Sections 6.1 and 6.2 we present the estimation results of two reference experiments. The setup of the reference experiments is without downscaling, that is the observation pixels coincide with the estimation pixels, and synthetic brightness observations are available once daily. The focus of Reference Experiment I is on the estimation of the initial condition, whereas Reference Experiment II is designed to examine the estimation of the model error. In another series of experiments, we address two topics which are crucial to the design of an operational soil moisture data assimilation system. First, we investigate the quality of the estimates as the horizontal resolution of the brightness data decreases (Section 6.3). Finally, we address the impact of the satellite repeat cycle on soil moisture estimation (Section 6.4). The computational requirements for the assimilation are discussed in Chapter 8.

6.1 Reference Experiment I

6.1.1 Experiment Design

For a synthetic experiment, we generate synthetic fields for the uncertain inputs. This allows us to specify fields of land surface variables which serve as the "true" states. These true states are derived with a single simulation or forward integration of the land surface model. Starting from a synthetically generated initial condition field, the land surface model is driven with the measured micro-meteorologic inputs and synthetically generated process

noise. The resulting fields are then used as the true fields to which we can compare the estimates and the prior trajectories. See Section 5.1 for a more detailed description of synthetic experiments.

We use normal distribution functions when generating the (synthetic) true fields of the uncertain inputs, namely the uncertain initial condition parameter and the model error. As described in Section 4.5 we use exponential correlation functions for the temporal and spatial covariances. In this first experiment, Reference Experiment I, we choose a relatively large uncertainty for the initial condition of the saturation. The initial condition fields were generated around a spatially uniform prior top node saturation of 0.5. Alternatively, we could have spun up the model by integrating it for an extended period of time in order to derive possibly more realistic and heterogeneous prior initial conditions. But since we are conducting a synthetic experiment, the difference in the prior mean values matters little and the most important feature is the variance of the uncertain inputs.

The variance of the transformed saturation parameter γ_g has been selected such that the initial condition distribution for the top node saturation is almost uniform between a minimum allowed saturation of 0.3 and a maximum allowed saturation of 0.75. For the given realization, the sample mean is 0.55 and the sample standard deviation is 0.1. A snapshot of the true initial top node saturation just two hours after the initial time is shown in Figure 6.3. The correlation length for the initial condition parameters is $50km$. The initial saturation profile is set to be hydrostatic, that is the gravity force is exactly balanced by a pressure gradient, that is a gradient in soil saturation.

Finally, we have fixed the initial condition for the soil temperature at the air temperature at the initial time, which is at 10:00h local time (CDT) in the morning. Trials have shown that varying or estimating the initial soil temperature does not alter the results. This is because the memory of the soil temperature is only on the order of a few hours. An incorrectly specified initial soil temperature therefore makes very little difference over the two-week assimilation window.

In contrast to the high uncertainty in the initial saturation, we choose a comparably small model error for this first experiment. We account for model errors in the top flux boundary conditions for the saturation and the soil temperature by choosing appropriate ones and zeros in the matrix P_ω. We also account for model error in the canopy energy balance (3.13). The other energy balance equations are treated as perfect models and the matrix P_ν is determined accordingly. The standard deviation for the top moisture flux condition is the equivalent of $10W/m^2$ in latent heat flux. Likewise, the standard deviation of the forcing terms in the force-restore equation and in the canopy energy balance equation is set to $10W/m^2$. Note that the uncertain top flux boundary condition for the saturation affects all nodes that are connected to the surface by plants' roots. Owing to the rather shallow depth of the domain, all nodes are typically connected to the surface through the vegetation.

For the horizontal correlation of the model error we choose a small value of $1.5km$. This implies that only model errors in directly neighboring pixels are effectively correlated. In this way we can keep the computational effort for the evaluation of the convolution integral in the model error update small (Section 8.1.2). The correlation time for all model errors is 10 hours.

Figure 6.1 shows the area average true states for Reference Experiment I. The top panel shows the area average soil saturation at four out of the seven nodes. There is a general trend towards drier conditions across the domain over the two weeks of the experiment.

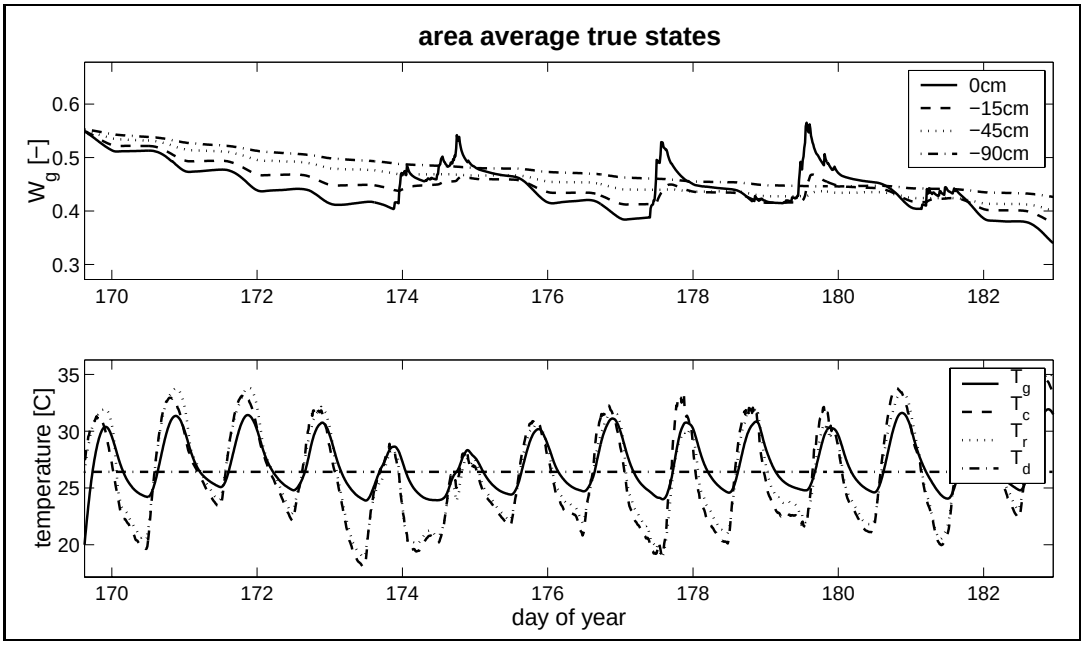

Figure 6.1: Area average true states for Reference Experiment I. The top panel shows the area average saturation at four out of seven layers. The precipitation events shown in Figure 5.2 are clearly discernible. The lower panel shows the area average soil temperature T_g, the canopy temperature T_c, and the deep soil temperature T_d together with the observed area average air temperature T_r. The deep soil temperature is derived by temporally averaging the measured air temperature as part of the micro-meteorologic inputs (Section 5.3.2). See Figure 5.2 for the area average micro-meteorologic inputs.

This drying trend is related to the fact that the southern half of the domain experienced only very little rainfall (Section 5.2). The lower panel shows the area average true soil and canopy temperatures. For comparison, the observed area average air temperature and the deep soil temperature are also shown. The deep soil temperature is specified as the time average of the observed air temperature (Section 5.3.2). As expected, the amplitude of the canopy temperature is larger than the amplitude of the soil temperature. There is also a short time lag between the canopy temperature and the soil temperature. The temperature of the canopy air space (not shown) tracks the canopy temperature very closely.

For the synthetic experiment, we neglect the interception water storage by setting κ^{W_c} and therefore the capacity of the interception water reservoir $W_c^{\max}$ to zero. The interception process is highly intermittent and interception happens on a time scale much shorter than the typical scale for the other state variables. Moreover, interception is almost by definition a threshold process and therefore not differentiable. For these reasons, the interception water is very difficult to estimate stably. Since the interception reservoir for pasture and crops is relatively small, neglecting the interception introduces only a small error. However, we feel that more research in this direction is necessary.

From the true states, synthetic L-band ($1.4GHz$) brightness data are derived on the scale of $5km$, which is also the scale of the estimation pixels. In this case the observation pixels coincide with the estimation pixels, and we refer to this scenario as a setup without downscaling. For twelve days, brightness images of the entire domain are generated once daily at 9:45h local time (CDT) in the morning. To each scalar true brightness temperature we add a spatially and temporally uncorrelated measurement error with a standard deviation of $5K$. Note that the absence of spatial correlation is not a constraint imposed by the algorithm. Uncorrelated measurement errors are attractive for a first application because they make coding easier and more transparent. In the future, spatially correlated measurement errors can be introduced rather straightforwardly. With 12 images and 512 brightness pixels per image, the total number of scalar data in the experiment is $12 \cdot 512 = 6144$.

Figure 6.2 summarizes the temporal setup of the reference experiments and other experiments which are discussed later. For the reference experiments, we use the entire two-week period as the assimilation window and measurements from all twelve observation times. Note that for some of the pixels some observations are taken during rain events. While it is raining, the water film covering the vegetation and the soil defies taking accurate measurements of the passive L-band microwave radiation. In an operational setting, a quality control routine must screen such faulty data. In this context, the rainfall interception model becomes important. For the synthetic experiments presented here, observations taken during rain events pose no particular problem.

6.1.2 Estimation of the True Fields

Figure 6.3 shows the true, the prior, and the estimated top node saturation for Reference Experiment I at six different times during the two-week period. In addition to the persistence of the initial condition, the evolution of the true top node saturation is governed by the precipitation inputs and the soil texture and land cover classes. This can best be observed in the prior fields because they start from a spatially uniform saturation. During the initial drydown (day 172.66) and during the second drydown (day 176.66), the soil texture and land cover properties lead to clearly visible streaky patterns which are most prominent in the northwestern corner of the experiment area. See Figure 5.3 for maps of the texture

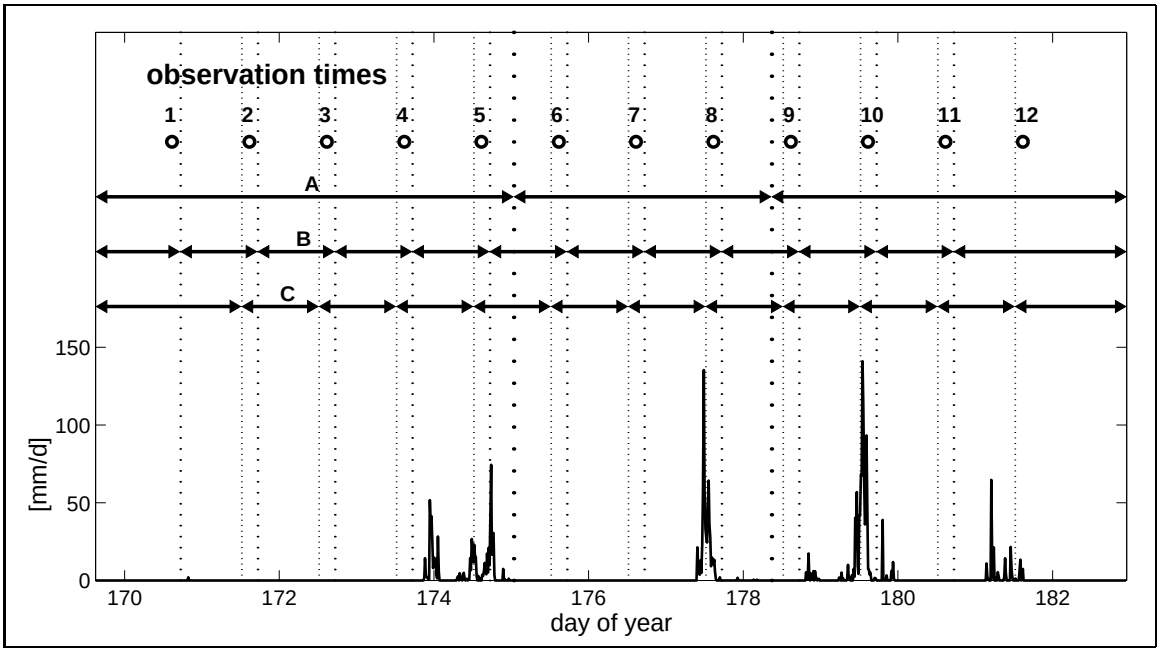

Figure 6.2: This Figure summarizes the temporal setup of the synthetic experiments. For reference, the area average precipitation time series is also shown. The experiment period covers two weeks from June 18, 1997 (day 169) to July 2, 1997 (day 183). There are up to twelve observation times at 9:45h local time (CDT) in the morning. In the reference experiments (Sections 6.1 and 6.2), all twelve observation times are used and there is just one assimilation (inversion) window which covers the entire two-week period. The same is true for the downscaling experiments (Section 6.3). The experiments on the brightness observation repeat cycle (Section 6.4) also use just one assimilation window. For the 3-day repeat cycle experiment, only the measurements at observation times 1, 4, 7, and 10 are assimilated. Likewise, only the measurements at observation times 1 and 7 are assimilated for the 6-day repeat cycle. Finally, the vertical dotted lines indicate the inversion time windows that are used in the three experiments with multiple assimilation windows (Section 7.1). In experiment A, we use three assimilation windows as indicated by the coarsely dotted vertical lines and the horizontal arrows. In experiments B and C, we use twelve assimilation windows for which the observation times are at the end and at the beginning of the window, respectively.

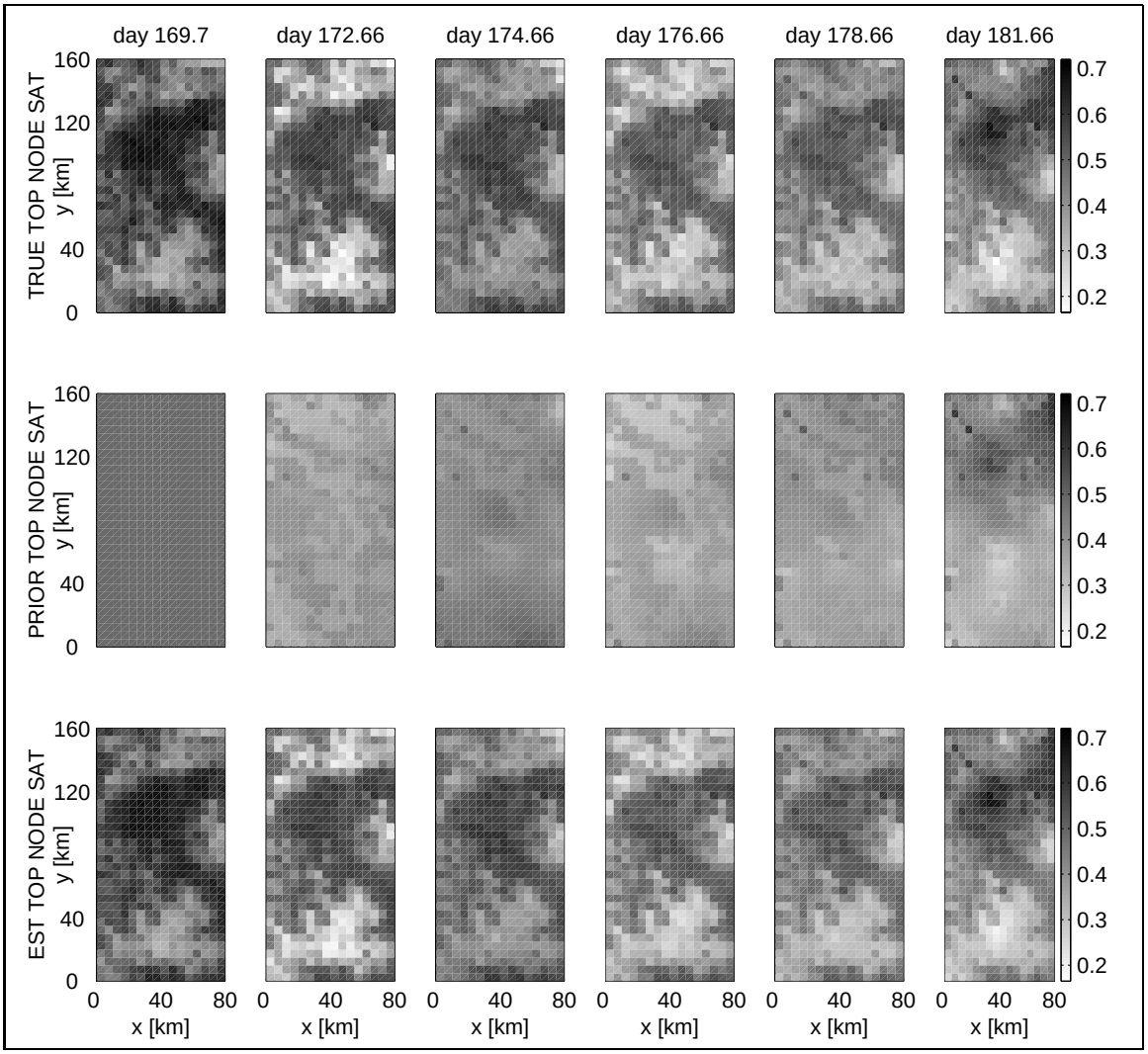

Figure 6.3: Top node saturation for Reference Experiment I. The first row shows the true top node saturation at six different times during the assimilation interval. In the second and third rows the prior and the estimate of the top node saturation, respectively, are depicted for the same six times.